Consciousness and the Binding Problem

WOLF SINGER

Max Planck Institute for Brain Research, Frankfurt/Main, Germany

ABSTRACT: It is proposed that phenomenal awareness, the ability to be aware of one’s sensations and feelings, emerges from the capacity of evolved brains to analyze their own cognitive processes by iterating and reapplying on themselves the very same cortical operations that they use for the interpretation of signals from the outer world. Search for the neuronal substrate of awareness therefore converges with the search for the cognitive mechanisms through which brains analyze their environment. The hypothesis is put forward that the mammalian brain generates continuously highly dynamic states that, when modulated by input signals, rapidly converge towards points of transient stability that correspond to the respective input constellation. It is proposed that these states are characterized by the dynamic binding of feature-specific cells into functionally coherent cell assemblies which as a whole represent the constellation of features defining a particular perceptual object. Arguments are presented that favor the notion that the cognitive operations supporting awareness consist of an iteration of such dynamic binding processes which then lead to the formation of higher-order assemblies that correspond to the contents of conscious awareness. Experimental data are reviewed relating to the questions of how assemblies are formed and which signatures define the relations among the responses of distributed neurons. It is argued that assemblies self-organize through reciprocal interactions of neurons coupled by reentrant loops and that the signature of relatedness consists of the transient synchronization of the discharges of the respective neurons. Evidence is presented that these synchronization phenomena depend on the same state variables as awareness: Both require for their manifestation activated brain states characterized by desynchronized EEGs. It is concluded that phenomenal awareness is amenable to neurobiological reductionism; but it is also proposed that self-consciousness requires a different explanatory approach because it emerges from the dialogue between different brains and hence has the quality of a cultural construct.

Keywords: Consciousness; Binding problem; Cerebral cortex; Synchronization; Visual perception; Awareness; Theory of mind; Social reality

The term consciousness has a number of different connotations ranging from awareness of one’s perceptions and sensations to self-awareness, the perception of oneself as an agent that is endowed with intentionality and free will. In this contribution I take the position that the first connotation of consciousness, phenomenal awareness, should in principle be tractable within neurobiological description systems because the problem can probably be reduced to the questions of how cognitive processes are

Address for correspondence: Max Planck Institute for Brain Research, Deutschordenstr. 46, 60528 Frankfurt am Main, Germany. Voice: ++49-(0)69-96769-218. singer@mpih-frankfurt.mpg.de
organized. The latter connotations, by contrast, transcend purely neurobiological descriptions because they have a social, a cultural, and a historical dimension.

AWARENESS

Brains capable of processing signals at a conscious level appear to have the ability to express the outcome of their distributed computational operations in a common format. They can run protocols of their own performance, both past and present, representing not only sensory and motor processes, but also the state of value-assigning systems. Thus, brains that have consciousness possess a meta-level at which their own internal states are subject to cognitive operations; they have what one might call an “inner eye” function. They can compare protocols of their own performance with incoming signals and derive from the outcome of these “internal deliberations” decisions for future acts. This allows them to respond with more flexibility to changing conditions than is the case for brains that lack consciousness and are confined to reacting to stimuli without the option of further reflection and internal deliberation. The implementation of such a reflexive meta-level has obvious adaptive functions and this may have contributed to the evolution of brains capable of being conscious of their own performance.

In order to run protocols of processes occurring within the brain, additional cognitive structures are required that analyze these processes. Thus, implementation of monitoring functions requires second-order processing of the computational results provided by first-order processes. The most likely substrate for such operations are cortical areas that have been added in the course of evolution and that treat the output of lower-order cortical areas in the same way as these treat input from the sensory periphery. The inner-eye function could thus be realized by a reflexive iteration of self-similar cortical processes. This interpretation is compatible with the neuroanatomical evidence that the phylogenetically more recent cortical areas are remote from primary sensory input and interact mainly either through the thalamus or directly through cortico-cortical connections with areas of lower order.

If these more recent monitoring structures have, in turn, access to the motor system—and available evidence indicates that this is the case—brains endowed with such monitoring functions would in addition have the possibility to signal to other organisms the result of the internal monitoring. Through mimickings, gestures, vocalizations and, in humans, language also, such brains could signal to others what their perceptions, intentions, value-assignments and plans of action are. Since such information increases dramatically the predictability of future actions of the respective other, it is likely to have an important function in the stabilization of labor-sharing societies, yet another adaptive function of consciousness that could have favored its evolution.

Two arguments, one based on evolution and the other on ontogeny, suggest that consciousness is a graded phenomenon whereby the gradations are correlated with the phylogenetic and ontogenetic differentiation of the cerebral cortex. The evolutionary argument is derived from the evidence that brains have evolved gradually, the most recent evolutionary changes being confined to an expansion of cerebral cortex and the apposition of new cortical areas. This suggests that consciousness evolved
as a consequence of cortical expansion and therefore is probably not an all-or-none phenomenon. The ontogenetic argument is based on the observation that the various manifestations of consciousness, from rudimentary awareness of sensations to the fully expressed self-consciousness of the adult, go in parallel with the gradual maturation of cerebral structures, in particular of the phylogenetically more recent cortical areas.

If one accepts the scenario that the aspect of consciousness that we address as phenomenal awareness results from an iteration of the same cognitive operations that support primary sensory processing, the explanatory gap reduces itself to the question of how the brain accomplishes its cognitive functions. If this question is answered with respect to primary sensory processes, the discovered strategies should be generalizable to the higher-order processes that serve the reanalysis of some of the brains’ own computational operations and assume the postulated inner-eye function.

SELF-CONSCIOUSNESS

Other aspects of consciousness such as self-awareness and the experience of individuality seem to require explanations that transcend purely neurobiological reductionism. It is my perception that the ontological status of these phenomena differs from that of the qualia of phenomenal awareness and that it is these aspects of consciousness that give rise to the hard problems in the philosophy of mind and provide the incentive for adopting dualistic positions. The most challenging phenomenon in this context is that we perceive ourselves as agents that are endowed with the freedom to decide, implying that the self is actually capable of controlling by will processes in the brain. We experience these aspects of consciousness as immaterial mental entities that are capable of influencing the neuronal processes required for the execution of actions and, hence, we perceive them as not derivable from the material processes in the brain.

I propose that these latter connotations of consciousness differ from phenomenal awareness because they cannot solely be accounted for by the performance of individual brains, but require for their development interactions among brains. These interacting brains need to be sufficiently differentiated to produce phenomenal awareness and to generate a theory of mind. In addition, they must be able to signal to one another and to comprehend that they are endowed with this capacity; they need to be able to enter dialogues of the kind that would say, for example, “I know that you know how I feel” or “I know that you know what my intentions are.” My proposal is that the concept of the “self” with all its subjective mental attributes, emerges from such dialogues among human beings, above all from the early interactions between caregivers and babies. Being told repeatedly: “Do this, otherwise you will be punished” or “leave this, because it is dangerous” is inevitably experienced as evidence that one is free to choose among one’s actions. The experience of individuality and responsibility, and as a consequence the intuition that one is endowed with free will would then have to be considered as a product of social interactions. The consequence is that the subjective attributes of consciousness would have the ontological status of social realities, of cultural constructs, and would, therefore, transcend pure neurobiological description systems that focus on individual brains.
The mechanisms that enable us to experience ourselves as endowed with mental capacities do, of course, reside in individual brains, but the contents of this experience are derived from social interactions. Why then should the experience of the self be so obviously different from other experiences that we derive from social interactions? Why do we experience the self as different from other social realities such as value systems or social rules. One explanation could be that the dialogue that leads to the experience of the self is already initiated during an early developmental stage, before episodic memory matures and begins to keep track of what the brain experiences. If so, there would be no conscious record of the processes that led to the experience of the self and the associated subjective connotations of consciousness. Because of this amnesia, these early experiences would lack causation, they would appear as timeless and detached from any real world context. Nevertheless, the contents of these early learning processes are bound to determine the way we act and experience ourselves. In consequence, the subjective connotations of consciousness, although they have been acquired by learning, would be perceived as having transcendent qualities that resist reductionistic explanations.

THE ORGANIZATION OF COGNITIVE PROCESSES

If the argument is valid that the internal monitoring functions are the result of an iteration of the same cognitive operations as the sensory processes which deal with signals conveyed by the sense organs, search for the neuronal substrate of phenomenal awareness converges with the search for the organization of cognitive processes in general. In the following paragraphs I shall, therefore, present hypotheses on the putative organization of cognitive operations in the mammalian brain.

Classical concepts of cognition are based on the assumption that the end result of a cognitive operation should be the generation of explicit neuronal representations of a particular content; in perception this content would be a distinct perceptual object that is characterized by a unique constellation of elementary features. Its neuronal representation is thought to consist of individual neurons that are tuned to particular constellations of input activity. Through their selective responses these neurons establish explicit representations of particular constellations of features. It is commonly held that the specificity of these neurons is brought about by selective convergence of input connections in hierarchically structured feed-forward architectures. This representational strategy allows for rapid processing and is ideally suited to the representation of frequently occurring stereotyped combinations of features; but this strategy is expensive in terms of the number of required neurons and not suited to cope with the virtually infinite diversity of possible constellations of features encountered in real-world objects. Such a representational strategy is also inappropriate for the encoding of syntactic structures and the hierarchical relations among elements of composite perceptual objects because it lacks systematicity (for review see Roelfsema et al.² Therefore, alternative concepts have been developed that emphasize more distributed dynamical processes that rely on self-organization. At the basis of these concepts is the assumption that neurons get associated into functionally coherent assemblies which as a whole stand for a particular content whereby each of the participating neurons is tuned to only a subset of the elementary features
of composite perceptual objects. This processing strategy is more economical with respect to neuron numbers because a particular neuron can, at different times, participate in different assemblies just as a particular feature can be part of many different perceptual objects. Moreover, this representational strategy is more flexible. It allows for the rapid de novo representation of constellations that have never been experienced before because there are virtually no limits to the dynamic association of neurons in ever-changing constellations, provided that the participating neurons are interconnected. The correlate of a particular percept would thus be a specific dynamic state of a large number of interacting but distributed neurons rather than the enhanced firing of a few highly specialized neurons at the top of a hierarchically organized feed-forward architecture.

The cognitive process required for the implementation of the inner-eye function has to cope with contents that are particularly unpredictable and rich in combinatorial complexity. The contents that need to be bound together are necessarily polymodal and change at the same pace as the contents of phenomenal awareness change. Thus, the cognitive processes supporting phenomenal awareness need to be especially flexible and able to deal with novelty, compositionality, and syntactic data structure; and this suggests that they, too, should be based on dynamic assembly formation rather than on explicit single-cell codes. If so, it follows that conditions favoring the formation of dynamically organized assemblies ought to be the same as those required for awareness to occur. As detailed below, brain states that are compatible with the manifestation of consciousness also favor the emergence of ordered spatio-temporal activity patterns that could serve as substrate for the formation of assemblies. By contrast, the response properties of individual neurons tend to differ only little in awake, sleeping, and anesthetized brains. It is unlikely, therefore, that the tuned responses of individual neurons are alone sufficient to support consciousness.

The following chapters will, therefore, focus on the question of whether there is any evidence for the existence of dynamically associated assemblies, and if so, what the electrophysiological manifestations of such assemblies might be. The hypothesis will be advanced that one signature of transiently stabilized assemblies is the synchronization of responses of participating neurons. Subsequently, data will be reviewed which suggest a correlation between perceptual processes and the occurrence of response synchronization, on the one hand, and between brain states favorable for the occurrence of awareness and the occurrence of synchronization, on the other.

**THE SIGNATURE OF ASSEMBLIES**

In assembly coding two important constraints need to be met. First, a selection mechanism is required that permits dynamic yet consistent association of neurons into distinct, functionally coherent assemblies. Second, responses of neurons that have been identified as groupable must get labeled so that they can be recognized by subsequent processing stages as belonging together. This is necessary in order to assure that responses, once they are bound together, are evaluated jointly as constituents of a coherent code and do not get confounded with responses of cells belonging to other, simultaneously formed assemblies that represent different contents. Numerous theoretical studies have addressed the question of how assem-
blies can self-organize on the basis of cooperative interactions within associative neuronal networks. Here I shall focus on the second problem of assembly coding: the question how responses of cells that have been grouped into an assembly can be tagged as related. An unambiguous signature of relatedness is absolutely crucial for assembly codes because, unlike the case in explicit single-cell codes, the meaning of responses changes with the context in which they are interpreted. Hence, in assembly coding false conjunctions are deleterious. Tagging responses as related is equivalent to assuring that they ignite selectively and conjointly corresponding assemblies at subsequent processing stages. This, in turn, can only be achieved by jointly raising the saliency of the selected responses, and here there are three options. First, non-grouped responses can be inhibited; second, the amplitude of the selected responses can be enhanced; and third, the selected cells can be made to discharge in precise temporal synchrony. All three mechanisms enhance the relative impact of the grouped responses at the next-higher processing level. Selecting responses by modulating discharge rates is common in labeled line coding, where a particular cell always signals the same content. This condition is usually met at low levels of processing, as in afferent and efferent pathways close to the respective sensory or effector organs. However, this strategy may not always be suited for the distinction of assemblies because it introduces ambiguities and reduces processing speed. Ambiguities could arise because discharge rates of feature selective cells vary over a wide range as a function of the match between stimulus and receptive field properties, and these modulations of response amplitude would not be distinguishable from those signaling the relatedness of responses. Processing speed would be reduced because rate-coded assemblies can only be identified after a sufficient number of spikes have been integrated to distinguish high from low rates. Therefore, they need to be maintained for some time in order to be distinguishable. This costs processing time, especially if several different assemblies need to be configured within the same matrix of neurons. Different assemblies cannot coexist in time if they share common subsets of neurons because it would be impossible to distinguish which responses belong to which assembly. Therefore, assemblies have to be multiplexed and configured in sequence. The rate at which different contents can be encoded does thus depend on the duration over which assemblies have to be maintained to be distinguishable. If defined exclusively by rate codes, assemblies need to have a lifetime of at least 50 to 100 ms, which slows down processing speed.

Both the ambiguities resulting from stimulus-related rate fluctuations and the temporal constraints can be overcome if the selection and labeling of responses is achieved through precise synchronization of individual discharges. Expressing the relatedness of responses by synchronization resolves the ambiguities resulting from stimulus-dependent rate fluctuations because synchronization can be modulated independently of rates. Response amplitudes could thus be reserved to signal how well particular features match the preferences of neurons, and synchronicity could be used in parallel to signal how these features are related. Defining assemblies by synchronization also accelerates the rate at which different assemblies can follow one another because the selected event is the individual spike or a brief burst of spikes; saliency is enhanced only for those discharges that are precisely synchronized and generate coincident synaptic potentials in target cells at the subsequent processing stage. The rate at which different assemblies can follow one another without getting
confounded is then limited only by the duration of the interval over which synaptic potentials summate effectively (for a detailed discussion, see Singer\textsuperscript{11}).

Another advantage of selecting responses by synchronization is that the timing of input events is preserved with high precision in the output activity of cells because synchronized input is transmitted with minimal latency jitter.\textsuperscript{12–14} This, in turn, can be exploited to preserve the signature of relatedness across processing stages, thus reducing further the risk of getting false conjunctions. Finally, synchronization enhances processing speed also by accelerating synaptic transmission per se because synchronized EPSPs trigger action potentials with minimal delay.

**PREREQUISITES FOR SELECTION BY SYNCHRONIZATION**

At the level of cellular mechanisms two prerequisites need to be fulfilled in order to exploit synchrony as a signature of relatedness: First, neurons must be able to act as coincidence detectors, that is, they must be particularly sensitive to coincident synaptic inputs. Second, mechanisms must exist which permit rapid and context-dependent temporal coordination of distributed discharge patterns.

The question of whether neurons in the central nervous system are capable of performing coincidence detection with the required precision is controversial since both theoretical arguments and simulation studies led to opposite conclusions.\textsuperscript{13–15} However, experimental evidence indicates clearly that neurons can evaluate temporal relations among incoming activity with sometimes surprising precision. In the auditory system coincidence detection is used to locate sound sources. Neurons in auditory nuclei of the brainstem evaluate the delays among incoming signals from the two ears with a precision in the submillisecond range (for review see Carr\textsuperscript{16}). Another example is the oscillatory responses of retinal ganglion cells that can be synchronized over large distances with close to zero phase lag.\textsuperscript{17} Because of the high frequency of these oscillations (up to 100 Hz), the neuronal mechanism responsible for synchronization must operate with time constants in the millisecond range. This time-modulated activity is reliably transmitted up to cortical neurons as indicated by cross-correlation analysis between retinal ganglion cells and cortical neurons.\textsuperscript{18} The implication is that neurons along the transmission chain must have operated with integration time constants not longer than a half cycle of the oscillation, and hence no more than 5 ms. The ability of cortical networks to handle temporally structured activity with high precision can also be inferred from the abundant evidence on the oscillatory patterning and synchronization of neuronal responses in the neocortex (reviewed in Singer and Gray\textsuperscript{19}). Such temporally coordinated discharge patterns can only emerge and stabilize if the temporal structure of activity is preserved during synaptic transmission and does not get dispersed and smeared too much by temporal integration. In the awake, performing brain the oscillatory patterning of cortical responses is typically in the gamma frequency range (30 to 60 Hz), and synchronization peaks often have a width at base in the range of 10 to 15 ms, indicating that temporal integration intervals should be on average no longer than 10 ms.

If synchronization is to play a role as signature of assemblies it must be possible to synchronize discharges rapidly because of the constraints set by processing speed.
Early simulation studies, which used harmonic oscillators rather than single spiking neurons, showed that it may indeed take a few cycles before synchronicity is established through phase locking. However, later simulations with spiking neurons revealed that networks of appropriately coupled units can undergo sudden state changes whereby the synchronization of discharges and their oscillatory patterning occurs promptly and virtually simultaneously (for review see Singer et al.8).

Very rapid synchronization has been observed recently in the visual cortex of cats. When neurons were activated by the onset of an appropriately oriented grating stimulus, their initial responses were already better synchronized than expected from mere stimulus locking. Comparison between actual response latencies and immediately preceding fluctuations of the local field potential revealed that the response latency shifted as a function of the polarity of the preceding field potential fluctuation. Because these fluctuations were not independent between the different recording sites, response latencies became synchronized. Thus, coordinated fluctuations of excitability act like a dynamic filter and cause a virtually instantaneous synchronization of the very first discharges of responses. Since the spatio-temporal patterns of these fluctuations reflect the architecture of intracortical association connections, grouping by synchronization can be extremely fast and can still occur as a function of the pre-wired associational dispositions of the cortical network.

Evidence suggests that an oscillatory patterning of responses may be instrumental for the internal synchronization of neurons, in particular when interactions comprise substantial conduction delays or occur across polysynaptic pathways. Experiments in slices support this conjecture, showing that subthreshold oscillatory modulation of the membrane potential is ideally suited to establish synchronization. In cells with oscillating membrane potential, responses can become delayed considerably, whereby the maximally possible delay interval depends on oscillation frequency and can amount to nearly the duration of one cycle. With such a mechanism, responses to temporally dispersed EPSPs can become synchronized within less than an oscillation cycle in cells exhibiting coherent fluctuations of their membrane potential.

**FUNCTIONAL CORRELATES OF RESPONSE SYNCHRONIZATION**

Following the discovery of stimulus-related response synchronization among neurons in the cat visual cortex, numerous experiments have been performed in search of a correlation between the occurrence of response synchronization and particular stimulus configurations. The prediction to be tested was that synchronization probability should reflect some of the Gestaltcriteria according to which the visual system groups related features during scene segmentation. Among the grouping criteria examined so far are continuity, vicinity, similarity, and colinearity in the orientation domain, and common fate in the motion domain (Refs. 9, 25–27 for the cat; Ref. 28 for the monkey). So far, the results of these investigations are compatible with the hypothesis that the probability of response synchronization reflects the Gestaltcriteria applied for perceptual grouping (Fig. 1). Stimulus-specific response synchronization has been found within and across different areas, and even between hemispheres (for review see Singer and Gray10) and, most importantly, none of these
synchronization phenomena were detectable by correlating successively recorded responses. This indicates that synchronization was not due to stimulus locking, but to internal dynamic coordination of spike timing. Thus, the observed temporal coherence among responses exceeds that expected from mere co-variation of event-related rate changes.

Studies involving lesions\textsuperscript{29,30} and developmental manipulations\textsuperscript{31,32} indicate that the interactions responsible for these dynamic synchronization phenomena are mediated to a substantial extent by reciprocal cortico-cortical connections. The criteria for perceptual grouping should then be reflected in the architecture of these connections, and this postulate agrees with the evidence that cortico-cortical connections preferentially link neurons with related feature preferences (for review see Schmidt et al.\textsuperscript{33}).

**RESPONSE SYNCHRONIZATION AND BEHAVIORAL STATES**

Most of the early experiments in search of synchronization phenomena have been performed in lightly anesthetized animals and it was important, therefore, to investigate whether response synchronization occurs also during states, where the EEG is actually desynchronized, as is characteristic for the awake, attentive brain. Evidence from cats and monkeys indicates that highly precise, internally generated synchrony is considerably more pronounced in the awake than in the anesthetized brain. Whenever tested, and data are available from the primary visual cortex of cats and monkeys, the motion-sensitive areas MT and MST in monkeys and infero-temporal cortex of monkeys, the synchronization phenomena were readily demonstrable and showed a dependence on stimulus configuration similar to the synchronization measured under anesthesia (for review see Singer et al.\textsuperscript{8}).

Of particular interest in this context is the recent finding that response synchronization is especially pronounced when the global EEG desynchronizes and when the animals are attentive. Stimulating the mesencephalic reticular formation in anesthetized animals leads to a transient desynchronization of the EEG, resembling the transition from slow-wave sleep to rapid-eye-movement sleep. Munk et al.\textsuperscript{34} (1996) and Herculano-Houzel et al.\textsuperscript{35} (1999) have recently shown that stimulus-specific synchronization of neuronal responses is drastically facilitated when the EEG is in a desynchronized rather than in a synchronized state.

Direct evidence for an attention-related facilitation of synchronization has been obtained from cats that had been trained to perform a visually triggered motor response.\textsuperscript{36} Simultaneous recordings from visual, association, somatosensory and motor areas revealed that the cortical areas involved in the execution of the task synchronized their activity, predominantly with zero phase-lag, as soon as the animals prepared themselves for the task and focused their attention on the relevant stimulus. Immediately after the appearance of the visual stimulus, synchronization increased further over the recorded areas, and these coordinated activation patterns were maintained until the task was completed. However, once the reward was available and the animals engaged in consumatory behavior, these coherent patterns collapsed and gave way to low-frequency oscillatory activity that did not exhibit any consistent phase relations. This close correspondence between performance of an attention-demanding visuo-motor task and the occurrence of inter-areal synchronization...
FIGURE 1. Stimulus-dependence of neuronal synchronization in area MT of the visual cortex of a macaque monkey carrying out a fixation task. Neuronal responses were obtained from two cell groups with different directional preferences. The figure shows cross-correlograms and peri-stimulus-time histograms for four different stimulation conditions. The small insets indicate the receptive field locations (1,2) with respect to the fixation point (F) and the directional preference of the neurons (small arrows). (A) A single moving stimulus bar, whose direction of motion was intermediate between the neurons’ preferences, led to a pronounced synchronization of the two cell groups, as indicated by the central maximum in the cross-correlogram. (B) Presentation of two stimuli moving in the respective preferred directions of cell group 1 and 2 abolishes synchronization. (C, D) Changing orientation and direction of motion by 15° or (D) using one of the bars from the configuration in (B) had little influence on synchronization. Scale bars for the peri-stimulus-time histograms correspond to 40 spikes/sec. The continuous line superimposed on the correlograms represents a damped cosine function that was fitted to the data to assess the significance of the correlogram modulation. (Modified from Kreiter and Singer.)
FIGURE 1. Continued.
tion suggests a functional role of the temporal patterning of widely distributed neuronal activity. One possibility is that the synchronization observed during the preparatory period reflects an entrainment of activity into a temporal pattern that is common to selected cortical areas in order to facilitate rapid temporal coordination of signals and fast convergence towards points of stability once the stimulus has become available. Attentional mechanisms could impose a coherent subthreshold modulation on neurons in cortical areas that need to participate in the execution of the anticipated task and thereby permit rapid synchronization of selected responses using the synchronizing mechanisms described above. According to this scenario, the attentional mechanisms would induce what one might call a state of expectancy in the respective cortical areas by imposing on them a specific, task-related dynamic activation pattern which then, once stimulus-driven input becomes available, acts like a dynamic filter that causes rapid synchronization of selected responses, thereby accomplishing the required grouping and binding of responses and in addition assuring rapid transmission.

PERCEPTION

In a series of visual experiments attempts have been made to find correlations between perceptual disturbances and abnormalities in neuronal synchronization.

Kittens were made strabismic shortly after eye opening, which results in an inability to group into a coherent percept signals generated by the two eyes. This inability is reflected by the failure of neurons driven by the two eyes to synchronize their responses even if these are evoked by a single object. A likely reason for the disruption of response synchronization is that the tangential intracortical connections between neurons driven by different eyes are lost during early development due to consistent de-correlation of the responses evoked from the two eyes.

A subgroup of the strabismic animals developed, in addition, a syndrome called strabismic amblyopia. Subjects suffering from strabismic amblyopia, and this is true for both animals and humans, have reduced visual acuity in the amblyopic eye. Moreover, they have difficulties in identifying figures if these are embedded in a contour-rich background, suggesting problems with perceptual grouping. In these amblyopic animals there was a close correlation between perceptual deficits and abnormalities in response synchronization, but there was no evidence for abnormalities of the response properties of individual cells. Quite unexpectedly, neurons in the primary visual cortex responded equally well to visual stimuli irrespective of whether these were shown to the normal or to the amblyopic eye. Thus, neurons driven by the amblyopic eye continued to respond vigorously to gratings that the animals had not been able to resolve with this eye during previous behavioral testing. Responses mediated by the normal and the amblyopic eye showed no difference in their amplitude or feature specific tuning irrespective of the spatial frequency of the applied test gratings. The only significant difference was the reduced ability of neurons driven by the amblyopic eye to synchronize their responses, and this difference was particularly pronounced when the spatial frequency of the grating stimuli approached the range beyond which the animals had no longer been able to resolve gratings with the amblyopic eye.
In primary visual cortex, there is, thus, a close correlation between a specific perceptual deficit and alterations in synchronization; and this relation could very well be causal. If synchronization is used to group responses together and to label them for further joint processing, then one expects that disturbances in synchronization should lead to disturbances in perceptual grouping such as occur for interocular binding operations in all strabismic animals and for monocular grouping operations in animals suffering from strabismic amblyopia. Since reduced synchronization is likely to reduce the saliency of responses conveyed by the amblyopic eye, it can also account for the fact that the amblyopic eye consistently loses in interocular competition when both eyes are open. Here, then, is a clear case where the firing of neurons in a cortical area does not correspond to perception, suggesting that the firing of individual neurons is only a necessary but not a sufficient condition to support perception. Additional and, in this case, indispensable information appears to be conveyed by the precise temporal relations among the discharges of simultaneously active neurons.

Another close correlation between response synchronization and perception and a remarkable dissociation between responses of individual neurons and perception has been found in experiments on binocular rivalry. When the two eyes are presented with patterns that cannot be fused into a single coherent percept, the two patterns are perceived in alternation rather than as a superposition of their components. This implies that there is a central gating mechanism which selects in alternation the signals arriving from the two eyes for further processing. Interocular rivalry is thus a suitable paradigm to investigate the neuronal correlates of dynamic response selection, a process closely related to the formation of assemblies.

This paradigm has been applied to investigate how neuronal responses that are selected and perceived differ from those that are suppressed and excluded from supporting perception. Multunit and field potential responses were recorded with chronically implanted electrodes from up to 30 sites in cat primary visual cortex while the animals were exposed to rivalrous stimulation conditions. Because the animal performs tracking eye movements only for the pattern that is actually perceived, it was possible to infer from the optokinetic tracking response which of the two eyes is selected. The outcome of these experiments was surprising as it turned out that the discharge rate of neurons in primary visual cortex failed to reflect the alternating suppression and selection of input from the two eyes. A close and highly significant correlation existed, however, between changes in the strength of response synchronization and the outcome of rivalry. Cells mediating responses of the eye that won in interocular competition increased the synchronicity of their responses upon introduction of the rivalrous stimulus, while the reverse was true for cells driven by the eye that became suppressed. Thus, in this particular case of competition, selection of responses for further processing appears to be achieved by raising their saliency through synchronization rather than by enhancing discharge frequency. Likewise, suppression is not achieved by inhibiting responses but by desynchronization (Fig. 2). In other terms one could say that only activity exhibiting a sufficient degree of temporal coordination has access to conscious awareness.

As in the amblyopic animals, there is thus a remarkable dissociation, at least in primary visual areas, between perception and the discharge rate of individual neurons. Cells whose responses are not perceived and are excluded from controlling behavior respond as vigorously as cells whose responses are perceived and support
behavior. This dissociation is particularly stringent in the case of rivalry because here responses to physically unchanged stimuli were recorded from the same neurons before and after introducing the rivalrous stimulus. Responses could be followed continuously while they passed from a condition where they were readily perceivable to a condition where they either continued to support perception in spite of rivalry or became excluded from perception. Another puzzling result of the rivalry
study is that responses that win the competition increase their synchronicity upon presentation of the rivalrous stimulus. This suggests the action of a mechanism that enhances the saliency of the selected responses by improving their synchronicity in order to protect them against the interference caused by the rivalrous stimulus.

These results provide direct support for the hypothesis that precise temporal relations between the discharges of spatially distributed neurons matter in cortical processing and that synchronization may be exploited to jointly raise the saliency of the responses selected for further processing. The important point here is that this selection can obviously be achieved without inhibiting the non-selected responses. Thus, in principle, it should be possible to select a second group of responses by synchronizing them independently of the first. The result would be two coexisting, but functionally distinct, assemblies at the same processing level. The example of rivalry also illustrates how synchronization and rate modulation depend on each other. The signals from the suppressed eye failed to induce tracking eye movements, indicating that eventually the vigorous but poorly synchronized responses in primary visual areas failed to drive the neurons responsible for the execution of eye movements. Direct evidence for the failure of neurons at subsequent processing stages to respond to the input from the respective suppressed eye has been obtained in behaving monkeys trained to indicate which of the two eyes they were actually using. This interdependence of discharge rates and synchrony provides the option to use both coding strategies in parallel in order to encode complementary information (see above).

A particularly close correlation between neuronal synchrony and perceptual grouping has recently been observed in experiments with plaid stimuli. These stimuli are well suited for the study of dynamic binding mechanisms because minor changes of the stimulus cause a binary switch in perceptual grouping. Two superimposed gratings moving in different directions (plaid stimuli) may be perceived either as two sur-
faces, one being transparent and sliding on top of the other (component motion), or as a single surface, consisting of crossed bars, that moves in a direction intermediate to the component vectors (pattern motion). Which percept dominates depends on the luminance of grating intersections because this variable defines the degree of transparency. Component (pattern) motion is perceived when luminance conditions are compatible (incompatible) with transparency (Fig. 3A). Here is a case where local changes in stimulus properties cause global changes in perceptual grouping. In the case of component motion, responses evoked by the two gratings must be segregated and only responses evoked by the contours of the same grating must be grouped to represent the two surfaces; in the case of pattern motion, responses to all contours must be bound together to represent a single surface. If this grouping of responses is initiated by selective synchronization, three predictions must hold (Fig. 3B): First, neurons that prefer the direction of motion of one of the two gratings and have colinearly aligned receptive fields should always synchronize their responses because they respond always to contours that belong to the same surface. Second, neurons that are tuned to the respective motion directions of the two gratings should synchronize their responses in case of pattern motion because they then respond to contours of the same surface but they should not synchronize in case of component motion because their responses are then evoked by contours belonging to different surfaces. Third, neurons preferring the direction of pattern motion should also synchronize only in the pattern and not in the component motion condition.

An important aspect of these predictions is that the expected changes in synchrony differ for different cell pairs, depending on the configuration of their receptive fields. Thus, when searching for relations between synchrony and cognitive functions, it is not only crucial to identify the processing stage where one assumes a particular binding function to be accomplished, but also to select the appropriate cell pairs. Averaging data across cell pairs with different receptive field configuration can mask dynamic changes in synchrony and is likely to reveal only the static anisotropies in the network of synchronizing connections. Such a problem may have contributed to the negative results of a recent study which failed to show a relation between perceptual grouping and internal synchronization in monkey striate cortex.44

In the case of the plaid stimuli, predictions were tested with multielectrode recordings from areas 18 and PMLS of the visual cortex of lightly anesthetized cats after we had confirmed with eye-movement recordings in awake cats that the ani-

FIGURE 3. (A) Two superimposed gratings that differ in orientation and drift in different directions are perceived either as two independently moving gratings (component motion) or as a single pattern drifting in the intermediate direction (pattern motion), depending on whether the luminance conditions at the intersections are compatible with transparency. (B) Predictions on the synchronization behavior of neurons as a function of their receptive field configuration (left) and stimulation conditions (right). (C) Changes in synchronization behavior of two neurons recorded simultaneously from areas 18 and PMLS that were activated with a plaid stimulus under component (upper graph) and pattern motion conditions (lower graph). The two neurons preferred gratings with orthogonal orientation (see receptive field configuration [top] and tuning curves obtained with component and pattern, respectively) and synchronized their responses only when activated with the pattern stimulus (compare cross-correlograms on the right). (Reproduced by the courtesy of Miguel Castelo-Branco and Sergio Neuenschwander).
FIGURE 3. See previous page for caption.
mals distinguished between component and pattern motion. Cross-correlation analysis of responses from cell pairs distributed either within or across areas 18 and PMLS confirmed all three predictions. Cells synchronized their activity if they responded to contours that are perceived as belonging to the same surface 45 (Fig. 3C). Analysis of the neurons’ discharge rate confirmed that most of the cells in these visual areas respond preferentially to the component gratings of the plaids (component-specific cells 46) and not to the pattern as a whole. However, in contrast to synchrony, variations in response amplitude failed to reflect the transition from component to pattern motion induced by transparency manipulation. Dynamic changes in synchronization could thus serve to encode in a context-dependent way the relations among the simultaneous responses to spatially superimposed contours and thereby bias their association with distinct surfaces. Future investigations have to clarify whether the populations of differentially synchronized neurons already serve as the final representation of the perceived surfaces or whether in case of pattern motion additional assemblies are formed. These would then have to consist of conjunction units tuned to the specific constellations of superimposed gratings, and their responses would have to bind together to signal that they code for the same surface.

In conclusion the data reviewed in this chapter indicate that evaluation of internally generated correlation patterns between responses of simultaneously recorded neurons permits the extraction of information about stimulus configurations, behavioral states and perception that cannot be obtained by analyzing the responses of individual neurons sequentially. The relevant variable containing this additional information is the rather precise synchronization of a fraction of the discharges constituting the respective responses. The data indicate further that responses containing synchronized epochs have a higher probability of being processed further and eventually to be perceived than do responses lacking such synchronized epochs, supporting the hypothesis that synchronization is used as a mechanism for response selection. Since synchronization necessarily involves at least two neurons, it inevitably raises simultaneously the saliency of more than one response. It is thus well suited to select subsets of responses for further joint processing, thereby defining the group of selected responses as related. Thus, synchronization fulfills the requirements postulated for a binding mechanism that selects a subset from a larger number of simultaneously active neurons and labels the responses of this subset in a way that favors joint processing at the subsequent processing stage. The evidence that synchronization probability reflects Gestalt criteria that guide perceptual grouping supports the hypothesis that synchronization serves as a binding mechanism in the context of assembly formation by jointly raising the saliency of selected subsets of responses. Transient synchronization of distributed responses might thus be considered as the distinguished state to which the assembly forming self-organizing process converges during sensory-motor processing.

**THE GENERALITY OF SYNCHRONICITY**

Studies in non-visual sensory modalities and in the motor system indicate that synchrony and oscillatory activity may actually be quite ubiquitous in the nervous system. Synchronization in the gamma frequency range occurs in the olfactory sys-
tem of various vertebrate and invertebrate species, where it has been related to the processing of odor information (for review see Laurent47). In the auditory cortex, synchronized gamma oscillations have been described both in humans (for review see Joliot et al.,48) and in animals.49,50 In the somatosensory system, synchronized oscillatory activity in this frequency range has recently been described both in the cat51 and in the monkey.52 Furthermore, synchronized oscillatory firing has been observed in other systems like the hippocampus53 and the frontal cortex.54

Similar evidence is available for the motor system where neural synchronization in the gamma frequency range has been discovered in cats,51 monkeys,52,55 and humans.56

Synchrony may also play a role in sensorimotor integration. In awake, behaving monkeys, task-dependent synchronization between units distributed across sensory and motor cortical areas and within motor cortex itself has been reported,52,55,57–59 Thus, synchrony also seems to play a role in visuomotor coordination and in the programming of motor responses. As proposed previously, it may be instrumental for the flexible channeling (binding) of sensory signals to particular motor programs.2

Synchronization also seems to play a role in the linkage between cortical assemblies and subcortical target structures such as the superior colliculus. This possibility is suggested by the existence of precise temporal relationships between the discharges of neurons in areas of the visual cortex and the superior colliculus.60 In these experiments, it could be shown that cortico-tectal interactions are strongly dependent on the temporal coherence of cortical activity. If cortical neurons engage in synchronous oscillatory activity either with partners within the same cortical area or with cells in other cortical areas, their impact on tectal cells is enhanced, indicating that tectal cells are driven more effectively by synchronous than by asynchronous cortical activity. This finding is consistent with the idea that the temporal organization of activity patterns plays an important role in defining the output of the cortex.

Taken together, available evidence suggests that comparable synchronization phenomena are found in a large number of different functional systems. Thus, it seems justified to generalize the results obtained in the visual cortex and to suggest that temporal coordination of discharges may be of general relevance for neural information processing. Importantly, there is increasing evidence that dynamic synchronization, in particular at frequencies in the gamma band, occurs also in the human brain. EEG studies have provided evidence for precise synchronization of activity in the γ-frequency range in human visual cortex that is related to perception and reflects the subjective coherence of stimulus arrangements,61–65

SYNCHRONICITY AS A GENERAL CODE FOR RELATEDNESS

If synchronicity serves as the signature of relatedness, synchronized responses should be interpreted as being related, irrespective of the cause of synchronization. Psychophysical evidence supports this conjecture. It indicates that synchronously presented stimuli are bound perceptually and interpreted as elements of the same figure with greater probability than asynchronously appearing texture elements (Refs. 66–70; but see Ref. 71). Hence, the synchronicity of responses imposed by simultaneously appearing texture elements (stimulus-locked synchronization) seems to be exploited for perceptual grouping. With respect to their strength and temporal preci-
sion, the externally induced and internally generated synchrony is virtually indistinguishable. Since the psychophysical results indicate that the former is interpreted as a signature of relatedness, it would be puzzling if this were not the case also for internally generated synchrony. Synchronization could thus serve as a general tag of relatedness, irrespective of whether it results from coincidence of external events or from internal grouping operations.

CONCLUSION

The hypothesis defended here is based on the following assumptions:

1. Phenomenal awareness emerges from an iteration of self-similar cognitive operations;
2. These are accomplished by cortical networks, and iteration is achieved by the addition of cortical areas of higher order that process the output of lower-order areas in the same way as these process their respective input;
3. In order to account for the required combinatorial flexibility these cognitive operations are likely to be based on the dynamic, self-organizing association of distributed neurons into functionally coherent assemblies and not solely on the generation of content-specific responses of individual specialized cells;
4. The binding mechanism that groups neurons into assemblies and labels their responses as related is the transient synchronization of discharges with a precision in the millisecond range; hence brief episodes of synchronized firing may be seen as a signature of the self-organizing dynamic process that leads to the formation of functionally coherent assemblies of spatially distributed neurons; and
5. The formation of such dynamically associated, synchronized cell assemblies requires activated brain states characterized by “desynchronized” EEG and is facilitated by attentional mechanisms.

The first assumption differs from the others because it is a conceptual premise which by itself does not provide any experimentally testable predictions, but each of the following subordinate assumptions leads to predictions about structural and functional features of brains capable of supporting the cognitive operations that give rise to phenomenal awareness. The data reviewed above illustrate that there is supportive experimental evidence for many of these features. However, we are lacking the proof that the observed features actually serve the function that our theories assign to them. This is the case not only for the more recently discovered functional properties such as the transient synchronization of distributed neuronal responses, but also for the rate-modulated discharges of individual neurons. As long as we have no complete theory on the structure of neuronal codes, it cannot be decided whether a sequence of discharges of a particular unit signifies that this unit participates in an assembly that lasts as long as the sequence of discharges or whether the unit participates in several different, successively organized assemblies, or whether it represents a content on its own. This uncertainty is due to the difficulty in identifying assemblies. Assemblies can only be identified if one succeeds in recording simultaneously from a sufficiently large fraction of neurons actually participating in a particular assembly. For reasons detailed elsewhere, this is technically very demanding, and therefore attempts to identify assemblies are still at the very beginning. Thus, if it is the
case that the cognitive processes required for consciousness involve the organization of highly distributed assemblies, we are a long way from the direct identification of the neuronal correlates of even the simplest forms of consciousness.

As long as analysis remains confined to the activity of individual neurons it will remain difficult to decide whether a recorded response is only a necessary or whether it is a sufficient condition for consciousness. Obviously, neurons need to discharge in order to convey information; if a group of neurons in a particular transmission chain stops responding, the content conveyed by that group of neurons cannot be perceived. Hence, correlations between perceptual awareness and cellular responses indicate only that the discharges of cells at a particular processing stage are necessary for a particular content to reach the level of awareness. In order to find out whether additional prerequisites have to be fulfilled, such as the binding of these responses into widely distributed assemblies, variables need to be determined that permit assessment of order parameters beyond the level of single units. This can only be achieved with recording techniques that disclose the spatio-temporal activation profile of large numbers of neurons. The fact that the most global of these methods, the EEG, differentiates rather reliably between brain states where consciousness is or is not possible favors the hypothesis that consciousness-supporting cognitive processes require coordination of activity well beyond the level of single-cell firing. Consciousness manifests itself only during brain states characterized by “desynchronized” EEGs. These states, in turn, favor the occurrence of gamma oscillations and long-distance synchronization of neuronal responses with a precision in the millisecond range. It seems not unreasonable, therefore, to pursue the hypothesis that the cognitive operations required for phenomenal awareness to manifest itself consist of the formation of large assemblies of distributed neurons whose signature of relatedness is the internally generated synchronicity of discharges.

REFERENCES


